

(11)Publication number : 09-237163  
(43)Date of publication of application : 09.09.1997

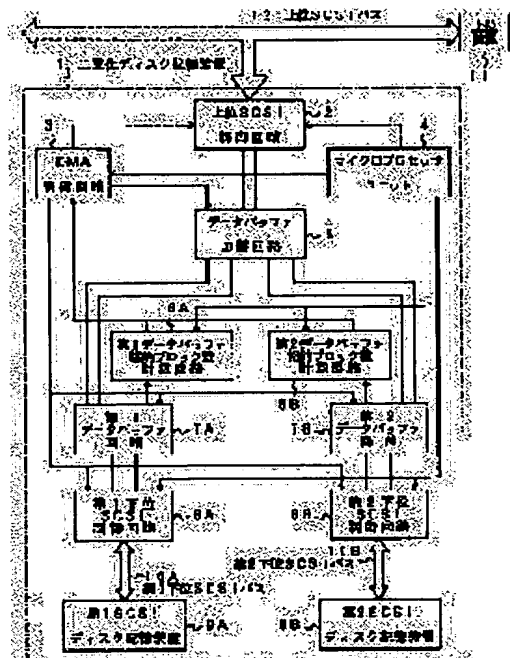
(51)Int.Cl.	G06F 3/06
	G06F 3/06
	G11B 19/02
	G11B 20/10

(21)Application number : 08-042253 (71)Applicant : NEC CORP  
(22)Date of filing : 29.02.1996 (72)Inventor : SAKAI TOSHIHIRO

**(57)Abstract:**

**PROBLEM TO BE SOLVED:** To prevent data transfer from being interrupted so as not to prolong a read processing time by a reason such as a case where a disk storage device selected at the time of reading allows reading to move to an exchange area within an instructed range and a case it becomes a rotation waiting state owing to a retry in a duplex disk storage device where the same data are written in the two disk storage devices.

**SOLUTION:** When a reading instruction is received from a host device 1, the first SCSI (small computer system interface) disk storage device 9A and the second SCSI disk storage device 9B respectively store independently read data in a first data buffer circuit 7A and the second data buffer circuit 7B, and a means is provided for counting the number of blocks stored in both circuits respectively the number of the blocks which are stored is reported to a DMA control circuit 3 at the point of time when storage is completed for each block. The DMA control circuit 3 permits data of the data buffer circuit stored with data earlier to be transferred red to the host device 11 through a high-order SCSI bus 12.



[Date of request for examination] 29.02.1996

[Date of sending the examiner's decision of rejection] 06.05.1998

[Kind of final disposal of application other than the examiner's decision of rejection or application converted registration]

[Date of final disposal for application]

[Patent number]

[Date of registration]

[Number of appeal against examiner's decision of rejection]

[Date of requesting appeal against examiner's decision of rejection]

[Date of extinction of right]

Copyright (C); 1998,2003 Japan Patent Office

(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開平9-237163

(43) 公開日 平成9年(1997)9月9日

(51) Int.Cl. <sup>6</sup>	識別記号	庁内整理番号	F I	技術表示箇所
G 0 6 F 3/06	3 0 4		G 0 6 F 3/06	3 0 4 E
	3 0 1			3 0 1 T
G 1 1 B 19/02	5 0 1		G 1 1 B 19/02	5 0 1 F
20/10		7736-5D	20/10	D

審査請求 有 請求項の数4 O L (全 6 頁)

(21) 出願番号 特願平8-42253

(22) 出願日 平成8年(1996)2月29日

(71) 出願人 000004237

日本電気株式会社

東京都港区芝五丁目7番1号

(72) 発明者 酒井 智弘

東京都港区芝五丁目7番1号 日本電気株式会社内

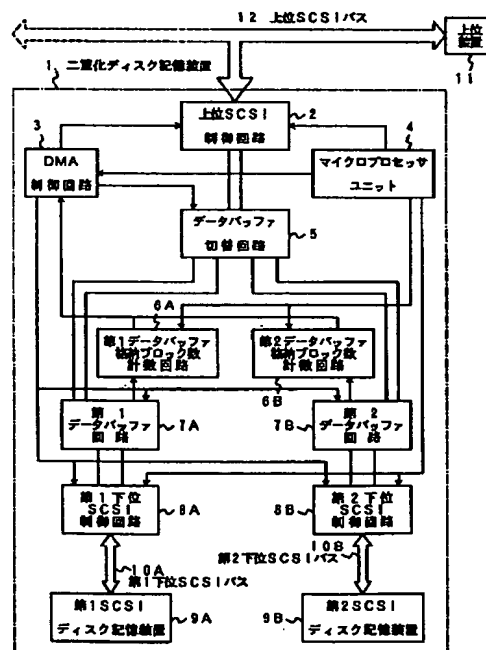
(74) 代理人 弁理士 京本 直樹 (外2名)

(54) 【発明の名称】 二重化ディスク記憶装置

(57) 【要約】

【課題】 2台のディスク記憶装置に同一のデータを書き込む二重化ディスク記憶装置において、読み出し時に選択されたディスク記憶装置が読み出しを指示された範囲内で交替領域へ移動を行う場合やリトライの為回転待ちの状態になるなどの理由によりデータ転送が中断してリード処理時間が延びることを防止する。

【解決手段】 上位装置11からの読み出し命令を受けると第1SCSIディスク記憶装置9Aと第2SCSIディスク記憶装置9Bとが独立に読み出したデータをそれぞれ第1データバッファ回路7Aと第2データバッファ回路7Bに格納し、両方の格納できたブロック数をそれぞれ計数する手段を持ち、ブロック毎に格納が完了した時点で格納したブロック数をDMA制御回路3に報告する。DMA制御回路3は早く格納し終えている方のデータバッファ回路のデータを上位SCSIバス12とが接続されるようにして上位装置11にデータを転送する。



**【特許請求の範囲】**

【請求項1】 第1のディスク記憶装置と第2のディスク記憶装置を有し、上位装置のデータ書き込み命令に対して前記第1のディスク記憶装置と前記第2のディスク記憶装置に同一のデータを書き込む二重化ディスク記憶装置において、

前記上位装置の読み出し命令に対して、

前記第1のディスク記憶装置と前記第2のディスク記憶装置から独立にデータを読み出し、

ブロック毎に早く読み出しが完了した方のディスク記憶装置のデータを前記上位装置に転送することを特徴とする二重化ディスク記憶装置。

【請求項2】 読み出し命令を受けた全ブロックの読み出しが早く完了した方の前記ディスク記憶装置の該全ブロックのデータ転送が終了した時点で、リード処理終了を前記上位装置に報告することを特徴とする請求項1記載の二重化ディスク記憶装置。

【請求項3】 第1のディスク記憶装置と第2のディスク記憶装置を有し、上位装置のデータ書き込み命令に対して前記第1のディスク記憶装置と前記第2のディスク記憶装置に同一のデータを書き込む二重化ディスク記憶装置において、

前記上位装置の読み出し命令に対して、第1のディスク記憶装置、第2のディスク記憶装置からそれぞれ独立に読み出したデータを格納する第1のデータ格納手段、第2のデータ格納手段と、

前記二重化ディスク記憶装置内のデータ転送の制御を行い、前記ブロック毎に前記第1のデータ格納手段と前記第2のデータ格納手段の格納状況を判定して、早く格納を完了している方のデータ格納手段と上位インタフェースとが接続されるよう接続切替手段に指示し、早く格納できたデータを前記上位装置へ転送するデータ転送制御手段と、

第1のデータ格納手段が格納できたブロック数を計数し、前記第1のデータ格納手段にブロック毎に格納が完了した時点で格納し終えたブロック数を前記データ転送制御手段に報告する第1格納ブロック計数手段と、

第2のデータ格納手段が格納できたブロック数を計数し、前記第2のデータ格納手段にブロック毎に格納が完了した時点で格納し終えたブロック数を前記データ転送制御手段に報告する第2格納ブロック計数手段と、

前記データ転送制御手段からの指示を受けて、前記第1のデータ格納手段及び前記第2のデータ格納手段と前記上位装置へ接続するための前記上位インタフェースとの接続を切り替える前記接続切替手段とを具備したことを特徴とする二重化ディスク記憶装置。

【請求項4】 読み出し命令を受けた全ブロックの読み出しが早く完了した方の前記ディスク記憶装置の該全ブロックのデータ転送が終了した時点で、前記リード処理終了を前記上位装置に報告することを特徴とする請求項

3記載の二重化ディスク記憶装置。

**【発明の詳細な説明】****【0001】**

【発明の属する技術分野】本発明は、二重化ディスク記憶装置の制御方式に関し、特に二重化ディスク記憶装置からのデータの読み出し処理の制御方式に関する。

**【0002】**

【従来の技術】ディスク記憶装置は、大容量のデータを記憶する装置であり、故障により記憶していたデータが消失してしまうことはコンピュータシステムにとって致命的問題となる。この問題を防ぐため、2台のディスク記憶装置に同一のデータを記憶することにより、1台のディスク記憶装置が故障してももう1台の正常なディスク記憶装置により正常運転を継続する、いわゆる二重化ディスク記憶装置が知られている。

【0003】従来は、二重化ディスク記憶装置においてデータの読み出しを行う場合、2台のディスク記憶装置をあらかじめ正ディスクと副ディスクに区別しておき正ディスクにのみリード命令を発行してデータを読み出す方式がとられていた。

【0004】又特開平3-259321号公報は、2台のディスク記憶装置ともにリード命令を発行して最初にデータ転送が可能になったディスクのデータを読み出してリード時間を短縮する技術を開示している。

【0005】又特開平3-253933号公報は二重化ディスク制御装置内のメモリに、二重化媒体毎の各種状態を格納する管理テーブルを備え、リード時にその管理テーブルを参照してリード処理ディスクを決定し、その状態によりアクセスを行うと共に、次のリードアクセスディスクを他方のディスクに変更することにより双方の二重化媒体から交互にリードする技術を開示している。

**【0006】**

【発明が解決しようとする課題】従来の二重化ディスク記憶装置からのデータの読み出し処理においては、あらかじめ決めておいた正ディスクのみから読み出したり、特開平3-259321号公報の場合は2台のうち早くデータ転送を開始した方のディスク記憶装置のデータを転送する方式のため、上位装置に転送されるデータは2台のディスク記憶装置のうちの1台のみから読み出したデータである。

【0007】このため、選択した方のディスク記憶装置において、読み出しを指示した範囲のブロックに交替割付処理されたブロックがあり交替領域へ移動を行う場合や、読み出しに失敗してリトライのため回転待ちの状態となる場合などの理由により、データ転送が中断する休止時間が発生しリード処理時間が大きく延びるという問題点がある。

【0008】又特開平3-253933号公報の場合は2台のディスクを交互に使う前提なのでリードを連続して何回も行う場合しか効果はあられず、独立した1回

のリードでは効果がないという問題点がある。さらに読み出しの途中で交替領域へ移動を行うときや、読み出しに失敗してリトライのため回転待ちの状態となるときはデータ転送が中断する休止時間が発生し、リード処理時間が大きく延びるという問題点がある点は特開平3-259321号公報の場合と同様である。

【0009】

【発明の目的】本発明の目的は、二重化ディスク記憶装置からのデータの読み出しの処理時間を短縮することにある。

【0010】

【課題を解決するための手段】本発明の二重化ディスク記憶装置は第1のディスク記憶装置と第2のディスク記憶装置を有し、上位装置のデータ書き込み命令に対して前記第1のディスク記憶装置と前記第2のディスク記憶装置に同一のデータを書き込む二重化ディスク記憶装置において、前記上位装置の読み出し命令に対して、前記第1のディスク記憶装置と前記第2のディスク記憶装置から独立にデータを読み出し、ブロック毎に早く読み出しが完了した方のディスク記憶装置のデータを前記上位装置に転送することを特徴とする。

【0011】本発明の二重化ディスク記憶装置は読み出し命令を受けた全ブロックの読み出しが早く完了した方の前記ディスク記憶装置の該全ブロックのデータ転送が終了した時点で、リード処理終了を前記上位装置に報告することを特徴とする。

【0012】本発明の二重化ディスク記憶装置は第1のディスク記憶装置と第2のディスク記憶装置を有し、上位装置のデータ書き込み命令に対して前記第1のディスク記憶装置と前記第2のディスク記憶装置に同一のデータを書き込む二重化ディスク記憶装置において、前記上位装置の読み出し命令に対して、第1のディスク記憶装置、第2のディスク記憶装置からそれぞれ独立に読み出したデータを格納する第1のデータ格納手段、第2のデータ格納手段と、前記二重化ディスク記憶装置内のデータ転送の制御を行い、前記ブロック毎に前記第1のデータ格納手段と前記第2のデータ格納手段の格納状況を判定して、早く格納を完了している方のデータ格納手段と上位インタフェースとが接続されるよう接続切替手段に指示し、早く格納できたデータを上位装置へ転送するデータ転送制御手段と、第1のデータ格納手段が格納できたブロック数を計数し、第1のデータ格納手段にブロック毎に格納が完了した時点で格納し終えたブロック数をデータ転送制御手段に報告する第1格納ブロック計数手段と、第2のデータ格納手段が格納できたブロック数を計数し、第2のデータ格納手段にブロック毎に格納が完了した時点で格納し終えたブロック数をデータ転送制御手段に報告する第2格納ブロック計数手段と、前記データ転送制御手段からの指示を受けて、前記第1のデータ格納手段及び前記第2のデータ格納手段と前記上位装置

へ接続するための上位インタフェースとの接続を切り替える前記接続切替手段とを具備したことを特徴とする。

【0013】データ転送制御手段により、第1のデータ格納手段と第2のデータ格納手段は、それぞれが接続されている第1と第2のディスク記憶装置から読み出したデータを格納する。第1と第2の格納ブロック数計数手段は、それぞれが対応する第1と第2のデータ格納手段に格納が完了した時点で格納し終えたブロック数をデータ転送制御手段に報告する。データ転送制御手段は、ブロック毎に第1と第2の格納手段の格納状況を判定して、早く格納し終えている方の格納手段と上位インタフェースとが接続されるよう接続切替手段に指示し、早く格納できたデータを上位装置に転送する。

【0014】このため、2台のディスク記憶装置から転送してくるデータのうち早く読み出せた方のデータを上位装置に転送し、早く読み出しを終えた方のディスク記憶装置の全ブロックの読み出し終了時点で上位装置へリード処理終了を報告することができる。

【0015】

【発明の実施の形態】次に、本発明について図面を参照して詳細に説明する。

【0016】図1は、本発明の実施形態を示すブロック図である。図1において、二重化ディスク記憶装置1は、上位SCSI (Small Computer Systems Interface) バス12を介して上位装置11と接続されている。二重化ディスク記憶装置1は、上位SCSIバス12に接続されターゲットとして動作する上位SCSI制御回路2を有する。二重化されたデータを保持する第1SCSIディスク記憶装置9A、第2SCSIディスク記憶装置9Bは、それぞれ第1下位SCSIバス10A、第2下位SCSIバス10Bを介して第1下位SCSI制御回路8A、第2下位SCSI制御回路8Bに接続される。これら第1下位SCSI制御回路8A、第2下位SCSI制御回路8Bはイニシエータとして動作し、ターゲットとして動作する第1SCSIディスク記憶装置9A、第2SCSIディスク記憶装置9Bを制御する。

【0017】第1データバッファ回路7A、第2データバッファ回路7Bは、それぞれ第1下位SCSI制御回路8A、第2下位SCSI制御回路8Bにデータバスにより接続され、さらにデータバッファ切替回路5に接続される。データバッファ切替回路5はさらに、データバスにより上位SCSI制御回路2と接続され、データバスの接続を第1データバッファ回路もしくは第2データバッファ回路に切り替える。図1中では二重化ディスク記憶装置1内のデータバスを二重線で表す。

【0018】DMA制御回路3は、上位SCSI制御回路2、第1下位SCSI制御回路8A、第2下位SCSI制御回路8B、第1データバッファ回路7A、第2データバッファ回路7B、データバッファ切替回路5とロ

ーカルバスにて接続されており、二重化ディスク記憶装置1内のデータ転送を制御する。第1データバッファ格納ブロック数計数回路6Aおよび第2データバッファ格納ブロック数計数回路6Bはそれぞれ、第1データバッファ回路7Aと第2データバッファ回路7Bに格納されているデータのブロック数を計数し、その情報をDMA制御回路3に報告する。

【0019】マイクロプロセッサユニット4は、上位SCSI制御回路2、第1下位SCSI制御回路8A、第2下位SCSI制御回路8B、DMA制御回路3、第1データバッファ格納ブロック数計数回路6A、第2データバッファ格納ブロック数計数回路6Bにローカルバスにて接続されており、二重化ディスク記憶装置1全体の動作を制御する。

【0020】次に、図2を併用して本発明の二重化ディスク記憶装置におけるデータの読み出し動作について詳細に説明する。

【0021】上位SCSI制御回路2が上位装置11からのリード・コマンドを受け取ると、第1下位SCSI制御回路8A、第2下位SCSI制御回路8Bはそれぞれに接続されている第1SCSIディスク記憶装置9A、第2SCSIディスク記憶装置9Bに対しリード・コマンドを発行してデータの読み出しを要求する。第2は、10ブロックのリードが要求された場合の具体例を示す。第1SCSIディスク記憶装置9A、第2SCSIディスク記憶装置9Bから読み出されたデータは、それぞれ第1下位SCSI制御回路8A、第2下位SCSI制御回路8Bを経て、第1データバッファ回路7A、第2データバッファ回路7Bに一時的に貯えられる。

【0022】このデータの受け取りは第1の系列（第1SCSIディスク記憶装置9A、第1下位SCSI制御回路8A、第1データバッファ回路7A）と第2の系列（第2SCSIディスク記憶装置9B、第2下位SCSI制御回路8B、第2データバッファ回路7B）とで互いに影響を及ぼすことなく独立に行われ、一方の系列のデータの受け取りが遅れてもそれとは無関係に他方の系列は先行してデータの受け取りを行う。

【0023】図2では、第1SCSIディスク記憶装置9Aの方が早くデータ転送を開始した場合を示している。第1データバッファ格納ブロック数計数回路6Aおよび第2データバッファ格納ブロック数計数回路6Bはそれぞれ、第1データバッファ回路7Aと第2データバッファ回路7Bが1ブロック分のデータを格納する毎にカウントアップして、第1データバッファ回路7Aと第2データバッファ回路7Bが格納したブロック数を計数する。このブロック数はマイクロプロセッサユニット4により新しい転送命令が出されることによりクリアされている。

【0024】DMA制御回路3は1ブロック毎に、第1データバッファ格納ブロック数計数回路6Aおよび第2

データバッファ格納ブロック数計数回路6Bが保持する格納ブロック数の情報を読みとり、第1データバッファ回路7Aと第2データバッファ回路7Bのうち該当するブロックのデータを早く格納し終えている方のデータを上位SCSI制御回路2に転送する。このときDMA制御回路3は、データバッファ切替回路5に指示をだしてデータバスを切り替えることにより、上位SCSI制御回路2と第2データバッファ回路7A、あるいは上位SCSI制御回路2と第2データバッファ回路7Bを接続する。

【0025】図2では、1ブロック目から5ブロック目までは第1データバッファ回路7Aの方が早くデータを格納しているため第1データバッファ回路7Aのデータが上位SCSI制御回路2に送られている。

【0026】図2では、第1SCSIディスク記憶装置9Aが交替領域への移動やリトライ処理を行った等の理由により6ブロック目のデータが転送されるまでに転送が休止してしまう場合を示しており、その間に1ブロック目のデータ転送の開始が遅かった第2SCSIディスク記憶装置9Bの方が6ブロック目以降のデータを早く第2データバッファ回路7Bに格納してDMA制御回路3の指示により上位SCSI制御回路2に送る場合を一例にあげている。上位SCSI制御回路2は第1データバッファ回路7A、あるいは第2データバッファ回路7Bから受け取ったデータを上位装置11に転送する。第1SCSIディスク記憶装置9Aと第2SCSIディスク記憶装置9Bのどちらか早い方が全ブロックのデータ転送を終了した時点で、DMA制御回路が第1データバッファ格納ブロック数計数回路6Aと第2データバッファ格納ブロック数計数回路6Bの内容からDMA制御回路3が終了と判断し、上位装置11に対しリード処理の終了を報告する。

【0027】図2では、1ブロック目から5ブロック目までは第1SCSIディスク記憶装置9Aから読み出したデータを、6ブロック目から10ブロック目までは第2SCSIディスク記憶装置9Bから読み出したデータを上位装置11に転送し、第2SCSIディスク記憶装置9Bのリード処理が終了した時点で上位装置11に対し終了報告を行っている。まだデータ転送の完了していない遅い方の系列は、継続して一時格納しても良いし、メッセージを発行してリード処理を中止しても良い。図2では、遅い方の第1SCSIディスク記憶装置9Aのデータ全10ブロックを第1データバッファ回路7Aに格納する場合を示している。

【0028】図2では図示していないが、第1系列と第2系列の切替は何度でも可能である。例えば転送中断がまず第1SCSIディスク記憶装置9Aで起こり、第2データバッファ回路7Bのブロックのデータが第1データバッファ回路7Aのそれに追いついた後に、第2データバッファ回路7Bのデータに切替えられ、その後、第

2 SCS Iディスク記憶装置9 Bのデータに転送中断が起きた場合は、第1 SCS Iディスク記憶装置9 Aのデータ転送が回復しており、かつ第2 SCS Iディスク記憶装置9 Bの転送中断が起きたブロックまで第1データバッファ回路7 Aにデータが転送されている場合には直ちに第1 SCS Iディスク記憶装置9 Aのデータに切り替える。

【0029】まだ転送中断が起きたブロックまで第1データバッファ回路7 Aにデータ転送が終わってない場合又は第1 SCS Iディスク記憶装置9 Aと第2 SCS Iディスク記憶装置9 Bが同時にデータ中断が起きた場合にはどちらか先に復帰して、中断したブロックにデータ転送が追いつき、更に継続して転送できる側のデータバッファ回路のデータが上位SCS I制御回路2に入り、上位装置11に転送されることになる。

【0030】

【発明の効果】以上説明したように、本発明による二重化ディスク記憶装置は、2台のディスク記憶装置から転送してくるデータのうち早く読み出せた方のデータを上位装置に転送するため、早く読み出しを終えた方のディスク記憶装置の読み出し終了時点で上位装置へ読み出し処理終了を報告することができ、交替領域への移動やリトライによる処理時間の増大を防ぐことができるという効果がある。

【0031】さらにリードを連続して何回も行う場合の

みならず、独立したリードを1回行う場合でも処理時間の増大を防げるという効果がある。

【図面の簡単な説明】

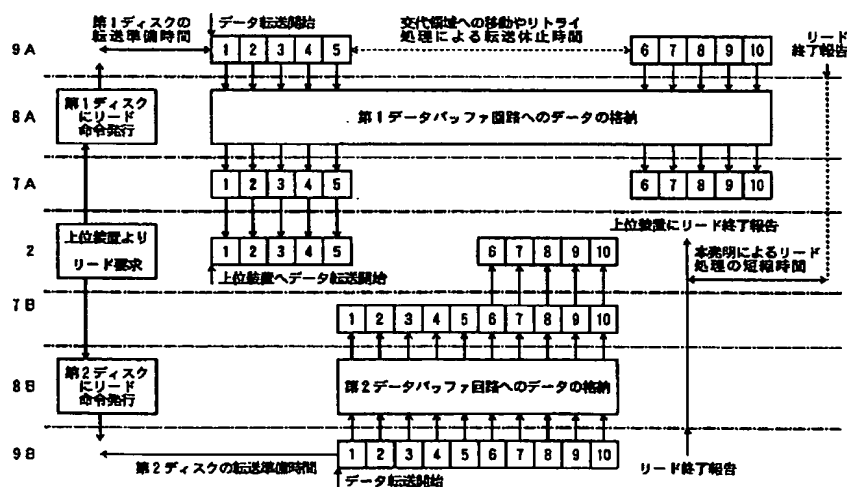
【図1】本発明の二重化ディスク記憶装置の一実施の形態を示すブロック図である。

【図2】本発明の二重化ディスク記憶装置の一実施の形態を示すシーケンス図である。

【符号の説明】

- 1 二重化ディスク記憶装置
- 2 上位SCS I制御回路
- 3 DMA制御回路
- 4 マイクロプロセッサユニット
- 5 データバッファ切替回路
- 6 A 第1データバッファ格納ブロック数計数回路
- 6 B 第2データバッファ格納ブロック数計数回路
- 7 A 第1データバッファ回路
- 7 B 第2データバッファ回路
- 8 A 第1下位SCS I制御回路
- 8 B 第2下位SCS I制御回路
- 9 A 第1SCS Iディスク装置
- 9 B 第2SCS Iディスク装置
- 10 A 第1下位SCS Iバス
- 10 B 第2下位SCS Iバス
- 11 上位装置
- 12 上位SCS Iバス

【図2】



【図1】

